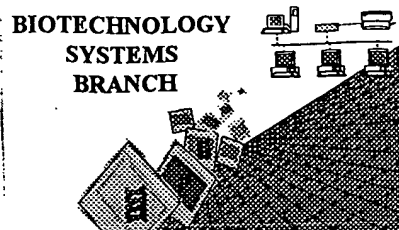# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number: *09/980,845*

Source: *PCT/09*

Date Processed by STIC: *5/21/2002*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE
    APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A
    NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
 PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
 PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
        http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is
a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the
United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission
   User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
   Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
           Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
   2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
   Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/980,845

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ✓ Variable Length

Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 ____ Skipped Sequences (OLD RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences (NEW RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 ____ Invalid <213> Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 ____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

PCT09

RAW SEQUENCE LISTING　　　　　　　DATE: 05/21/2002
PATENT APPLICATION: US/09/980,845　　TIME: 15:25:08

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05212002\I980845.raw

Does Not Comply
Corrected Diskette Needed

pp 1, 3-5

```
   3 <110> APPLICANT: Handfield, Martin
   4        Brady, Jeannine
   5        Progulske-Fox, Ann
   6        Hillman, Jeffrey D.
   8 <120> TITLE OF INVENTION: Microbial Polynucleotides Expressed During Infection of
   9        a Host
  11 <130> FILE REFERENCE: MBHB00-505B
C-->  13 <140> CURRENT APPLICATION NUMBER: US/09/980,845
C-->  14 <141> CURRENT FILING DATE: 2002-04-08
  16 <150> PRIOR APPLICATION NUMBER: 60/147,551
  17 <151> PRIOR FILING DATE: 1999-08-06
  19 <150> PRIOR APPLICATION NUMBER: PCT/US00/21340
  20 <151> PRIOR FILING DATE: 2000-08-04
  22 <160> NUMBER OF SEQ ID NOS: 20
  24 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

```
 777 <210> SEQ ID NO: 20
 778 <211> LENGTH: 54
 779 <212> TYPE: PRT
 780 <213> ORGANISM: Actinobacillus actinomycetemcomitans
 782 <400> SEQUENCE: 20
 783 Met Val Gly Lys Phe Ile Val Ile Glu Gly Leu Glu Gly Ala Gly Lys
 784  1               5                  10                  15
 786 Ser Thr Ala His Gln Cys Val Val Asp Thr Leu Lys Thr Leu Gly Val
 787             20                  25                  30
 789 Gly Glu Val Ile Ser Thr Arg Glu Pro Gly Gly Thr Pro Val Gly Gly
 790             35                  40                  45
 792 Lys Ala Thr Pro Ser His
 793          50
E-->  799 ②
```

delete

see following pages for more errors

```
120 <212> TYPE: DNA
121 <213> ORGANISM: Actinobacillus actinomycetemcomitans
123 <220> FEATURE:
124 <221> NAME/KEY: misc_feature
125 <222> LOCATION: (554)
126 <223> OTHER INFORMATION: N stands for any nucleotide.
128 <220> FEATURE:
129 <221> NAME/KEY: misc_feature
130 <222> LOCATION: (596)
131 <223> OTHER INFORMATION: N stands for any nucleotide.
133 <400> SEQUENCE: 3
134 gatcaaactg gtggcgcaag ggcagcgcgt agcaaattta cccgatattt tggtctatgc 60
135 gcgcgtcggc aacggcatgg tagggcgacg ccgtggttta aaccaagcca aagcggaatg 120
136 gcgcttattt aagctaaaac accatcttgg cattcaggga ttttttatccg ggctattcac 180
137 ttttgtcctg cgttccggtg ccagattatt gccgacatca ttactgaaaa acatctatca 240
138 aacctttta agaaaataac atgatgaaat taaactgtat tttaaaaata tccggaattt 300
139 ccaccgcact ttttctagcg ggttgttcct caaattcaag tgcgccgacg caatcctctg 360
140 agcaggcgaa ttctgttacg gctgtgaatc ccactgcggt gtacagtaag ccccgcactt 420
141 tggataactt caacgattat gtgaatttct taaaaggtaa agcagcggca gaaggcgttt 480
142 ctgccgacgt attgaatgca caaaataata ttaattatat tcaaaaatcc gtggatttgg 540
143 acgatcaaca agcnggcaga attcgcaagc gtgatccaaa tgccccgccg atcatnaatt 600
144 ccgaacggca cgaccaatta cttaaatcgt gtattaacca agaataaagt agacacggca 660
145 gaagcacgtt attgggaaca attgccgcag cttgaaaatg cttcaaagaa attcagcgta 720
146 ccgaaaaatt atctgttagc cttgtggggc atggagagta gctttggcta ttatcagggc 780
147 aattacgatg tgttatccac cttagccact cttgctttttg acggacgccg tgaagcctta 840
148 ttcagcaaag aattcatcgc cgccatgaaa atgctacagc gcgatc              886
151 <210> SEQ ID NO: 4
152 <211> LENGTH: 507
153 <212> TYPE: DNA
154 <213> ORGANISM: Actinobacillus actinomycetemcomitans
156 <220> FEATURE:
157 <221> NAME/KEY: misc_feature
158 <222> LOCATION: (4)
159 <223> OTHER INFORMATION: N stands for any nucleotide.
161 <220> FEATURE:
162 <221> NAME/KEY: misc_feature
163 <222> LOCATION: (9)
164 <223> OTHER INFORMATION: N stands for any nucleotide.
166 <220> FEATURE:
167 <221> NAME/KEY: misc_feature
168 <222> LOCATION: (21)
169 <223> OTHER INFORMATION: N stands for any nucleotide.
171 <220> FEATURE:
172 <221> NAME/KEY: misc_feature
173 <222> LOCATION: (23)
174 <223> OTHER INFORMATION: N stands for any nucleotide.
176 <220> FEATURE:
177 <221> NAME/KEY: misc_feature
178 <222> LOCATION: (29)
```

RAW SEQUENCE LISTING   DATE: 06/02/2002
PATENT APPLICATION: US/09/980,845   TIME: 19:10:38

          Input Set : A:\PTO.AMC.txt
          Output Set: N:\CRF3\05312002\I980845.raw

179 <223> OTHER INFORMATION: N stands for any nucleotide.
181 <220> FEATURE:
182 <221> NAME/KEY: misc_feature
183 <222> LOCATION: (32)
184 <223> OTHER INFORMATION: N stands for any nucleotide.
186 <220> FEATURE:
187 <221> NAME/KEY: misc_feature
188 <222> LOCATION: (35)..(36)
189 <223> OTHER INFORMATION: N stands for any nucleotide.
191 <220> FEATURE:
192 <221> NAME/KEY: misc_feature
193 <222> LOCATION: (39)
194 <223> OTHER INFORMATION: N stands for any nucleotide.
196 <220> FEATURE:
197 <221> NAME/KEY: misc_feature
198 <222> LOCATION: (42)
199 <223> OTHER INFORMATION: N stands for any nucleotide.
201 <220> FEATURE:
202 <221> NAME/KEY: misc_feature
203 <222> LOCATION: (45)
204 <223> OTHER INFORMATION: N stands for any nucleotide.
206 <220> FEATURE:
207 <221> NAME/KEY: misc_feature
208 <222> LOCATION: (49)
209 <223> OTHER INFORMATION: N stands for any nucleotide.
211 <220> FEATURE:
212 <221> NAME/KEY: misc_feature
213 <222> LOCATION: (52)
214 <223> OTHER INFORMATION: N stands for any nucleotide.
216 <220> FEATURE:
217 <221> NAME/KEY: misc_feature
218 <222> LOCATION: (58)
219 <223> OTHER INFORMATION: N stands for any nucleotide.
221 <220> FEATURE:
222 <221> NAME/KEY: misc_feature
223 <222> LOCATION: (61)..(62)
224 <223> OTHER INFORMATION: N stands for any nucleotide.
226 <220> FEATURE:
227 <221> NAME/KEY: misc_feature
228 <222> LOCATION: (65)
229 <223> OTHER INFORMATION: N stands for any nucleotide.
231 <220> FEATURE:
232 <221> NAME/KEY: misc_feature
233 <222> LOCATION: (69)
234 <223> OTHER INFORMATION: N stands for any nucleotide.
236 <220> FEATURE:
237 <221> NAME/KEY: misc_feature
238 <222> LOCATION: (73)
239 <223> OTHER INFORMATION: N stands for any polynucleotide.

"n" can only represent a single nucleotide (see item 5 on Erra Summary sheet) 6/2/02

```
     241 <220> FEATURE:
     242 <221> NAME/KEY: misc_feature
     243 <222> LOCATION: (97)
     244 <223> OTHER INFORMATION: N stands for any nucleotide.
     246 <220> FEATURE:
     247 <221> NAME/KEY: misc_feature
     248 <222> LOCATION: (102)
     249 <223> OTHER INFORMATION: N stands for any nucleotide.
     251 <220> FEATURE:
     252 <221> NAME/KEY: misc_feature
     253 <222> LOCATION: (138)
     254 <223> OTHER INFORMATION: N stands for any nucleotide.
     256 <220> FEATURE:
     257 <221> NAME/KEY: misc_feature
     258 <222> LOCATION: (457)
     259 <223> OTHER INFORMATION: N stands for any nucleotide.
     261 <220> FEATURE:
     262 <221> NAME/KEY: misc_feature
     263 <222> LOCATION: (459)
     264 <223> OTHER INFORMATION: N stands for any nucleotide.
     266 <220> FEATURE:
     267 <221> NAME/KEY: misc_feature
     268 <222> LOCATION: (467)
     269 <223> OTHER INFORMATION: N stands for any nucleotide.
     271 <400> SEQUENCE: 4
W--> 272 ttgntaccnt agccgctgac nanaactanc angcnntgna tnatntcgna tnattaanat 60
W--> 273 nngcnaggng cancagctta cctttgccga cggttcnctg tntgaaagcg ccattcgcaa 120
W--> 274 agtgccggtg gaggcggnga aaattcactc acttggtgcg gaaggcaatg atgtgggatt 180
     275 gaaagcccat catggcgggt ggataaagcg ttattttta tgtcggcaga tgcctttcct 240
     276 gcgttaaatg cgttattaga cgaaaatttt tcgtatcagg acacagcagt ttacggcgag 300
     277 aattttgtgg tttccgcgct gaatgaagat tccgtgtgtg tgggcgatat ttatcaaatc 360
     278 ggctcctgcg tggtggaggt gtcgcagccg cgtaaacctt gtgagcgctt atcgaaaaat 420
W--> 279 accaataatc cgaacacgca acaaaccgtg tacgctncng ctggtcnggc tggtatgtgc 480
     280 cggtggtacc ccaaggggga aattcaa                                   507
     283 <210> SEQ ID NO: 5
     284 <211> LENGTH: 1087
     285 <212> TYPE: DNA
     286 <213> ORGANISM: Actinobacillus actinomycetemcomitans
     288 <220> FEATURE:
     289 <221> NAME/KEY: misc_feature
     290 <222> LOCATION: (622)
     291 <223> OTHER INFORMATION: N stands for any nucleotide.
     293 <220> FEATURE:
     294 <221> NAME/KEY: misc_feature
     295 <222> LOCATION: (642)
     296 <223> OTHER INFORMATION: N stands for any nucleotide.
     298 <220> FEATURE:
     299 <221> NAME/KEY: misc_feature
     300 <222> LOCATION: (661)
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 06/02/2002
PATENT APPLICATION: US/09/980,845   TIME: 19:10:39

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF3\05312002\I980845.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 566,625,627,636,650,656,661,672,681,720,723
Seq#:3; N Pos. 554,596
Seq#:4; N Pos. 4,9,21,23,29,32,35,36,39,42,45,49,52,58,61,62,65,69,73,97
Seq#:4; N Pos. 102,138,457,459,467
Seq#:5; N Pos. 622,642,661,669,685,690,700
Seq#:6; N Pos. 609,614,651,665
Seq#:7; N Pos. 532,630,696,710,722,725
Seq#:8; N Pos. 538
Seq#:12; N Pos. 131,151,170,178,194,199,209
Seq#:18; Xaa Pos. 43,50,59,66,69